Journal of the Korean Statistical Society (2005), **34**: 4, pp 297–309

A NOTE ON PROTECTION OF PRIVACY IN RANDOMIZED RESPONSE DEVICES

AMITAVA SAHA[1]

ABSTRACT

We consider "efficiency versus privacy-protection" problem concerned with several well-known randomized response (RR) devices to estimate proportion of people bearing a stigmatizing characteristic in a community. The literature of RR on respondent's privacy protection discusses only about response specific jeopardy measures. We propose a measure of jeopardy that is independent of the RR offered by the interviewee and recommend it for using as a technical characteristic of the RR device. For ensuring better cooperation from the interviewees this new measure that depends only on the design parameters of the RR devices may be disclosed to the respondents before producing the RR by implementing the randomization device.

AMS 2000 subject classifications. Primary 62D05.
Keywords. Measure of jeopardy, protection of privacy, prior and posterior probabilities, unequal probability sampling.

1. INTRODUCTION

Randomized response (RR) technique was introduced by Warner (1965) as a tool for reducing answer bias and answer refusals and for protecting the respondent's privacy while collecting information on sensitive or stigmatizing features in estimating the proportion of people in a community bearing them. Different types of RR devices developed and studied by Warner (1965), Chaudhuri and Mukerjee (1988), Horvitz *et al.* (1967) and Greenberg *et al.* (1969), Mangat (1992, 1994) to name a few among many others give various degrees of protection to the respondent's privacy ensuring different levels of accuracies of the estimators.

Received June 2005; accepted December 2005.
[1]Directorate General of Mines Safety, Dhanbad, Jharkhand - 826001, India (e-mail : saha_amitava@hotmail.com)

An ideal RR device is one that allows a sufficient degree of protection of privacy of the respondents and at the same time produces good estimates of the population parameters. But, in general, greater protection of privacy is achievable at the cost of greater variance in estimation and thus protection of privacy and efficient estimation generally move in opposite directions.

Lanke (1976), Leysieffer and Warner (1976), Anderson (1976, 1977) and Nayak (1994) among many others discuss protection of the respondent's privacy in efficient estimation of parameters and studied the concept of 'respondent jeopardy'. But their work is exclusively confined to simple random sampling with replacement (SRSWR). Chaudhuri (2001a, b) extended some of the above-noted RR procedures to more complex survey situations where varying probability sampling without replacement is adopted.

The existing literature of RR on protection of interviewee's privacy discusses only about response specific jeopardy measures. That is, it discusses about the risk of revealing a respondent's status of bearing the stigmatizing characteristic only after the respondent provides the randomized response by applying the supplied randomization device. But as the measure of jeopardy is a quantified version of the risk of revelation of an interviewee's status regarding possession of a sensitive attribute, for the sake of ensuring better respondent cooperation, it is necessary to develop a response independent measure of jeopardy. This measure could be regarded as the measure of jeopardy of the RR device and may be made known to the respondents before they agree to use the RR device for providing the RR. Here we attempt to develop a technical characteristic of RR devices in terms of respondent's jeopardy, which is, of course, independent of the RR provided by the respondent and note that with increasing protection in privacy the efficiency of the estimator of the population parameter exhibits a declining trend.

In section 2 we briefly describe four RR-devices in complex survey situations where the samples are chosen with varying selection probabilities and in section 3 we present the measures of privacy-protection and the measures of jeopardy for them. Simulation-based numerical examples along with our comments are given in section 4.

2. EXAMPLES OF RR PROCEDURES IN COMPLEX SURVEYS

Let $U = (1, \cdots\cdots\cdots, N)$ be a finite population of N individuals with y_i as the value of a variable of interest on the ith individual such that

$$y_i = \begin{cases} 1 \text{ if the } i\text{th individual bears a stigmatizing characteristic } A \\ 0 \text{ if } i \text{ bears the complementary characteristic } A^C. \end{cases}$$

Our problem is to unbiasedly estimate $\pi_A = (\sum_{i=1}^N y_i)/N = Y/N$, writing, $Y = \sum_{i=1}^N y_i$ and assuming N to be known, *i.e.*, the proportion of individuals in the population bearing the stigmatizing feature A on choosing a sample, say, s of n individuals from U according to a sampling design p.

2.1. Warner's (1965) RR scheme

Each sampled individual is provided with a pack of cards marked A and A^C respectively in proportions $p : 1-p, 0 < p < 1$ and is requested to draw a card from this pack unnoticed by the interviewer and to report '1' or '0' according as the drawn card type matches his/her true A or A^C feature. Then writing

$$I_i = \begin{cases} 1 \text{ if the } i\text{th individual reports a 'match'} \\ 0 \text{ else,} \end{cases}$$

it follows that

$$P(I_i = 1) = py_i + (1-p)(1-y_i).$$

Now denoting E_R and V_R as the operators for expectation and variance respectively with respect to any RR scheme it is easy to check that

$$\begin{aligned} E_R(I_i) &= py_i + (1-p)(1-y_i) \\ V_R(I_i) &= p(1-p), \forall i \epsilon U, \end{aligned}$$

and defining $r_i = (I_i - (1-p))/(2p-1)$, $p \neq 1/2$ it follows that

$$E_R(r_i) = y_i \text{ and } V_R(r_i) = \frac{p(1-p)}{(2p-1)^2} = V_i, \text{ say, } \forall i \epsilon U.$$

2.2. Chaudhuri and Mukerjee's (1988) modification of Warner's scheme

Chaudhuri and Mukerjee (1988) proposed a modification of Warner's device. According to the modified scheme, each sampled individual is supposed to divulge

his/her true A or A^C character with a probability T_a or T_b. Here both A and A^C may be stigmatizing or only one of these may be so.
Now writing

$$I_i = \begin{cases} 1 \text{ if the characteristic } A \text{ or } A^C \text{, respectively} \\ \quad \text{matches the query about bearing } A \text{ or } A^C \\ 0 \text{ else,} \end{cases}$$

it follows that

$$P(I_i = 1) = T_a y_i + T_b(1 - y_i) = E_R(I_i).$$

Thus for $r_i = (I_i - T_b)/(T_a - T_b), T_a \neq T_b$, $E_R(r_i) = y_i$
and

$$\begin{aligned} V_i = V_R(r_i) &= \frac{E_R(I_i)[1 - E_R(I_i)]}{(T_a - T_b)^2} = \frac{1}{(T_a - T_b)^2}[T_b(1 - T_b) + (T_a - T_b)(1 - T_a - T_b)y_i] \\ &= \begin{cases} \frac{T_a(1-T_a)}{(T_a-T_b)^2} \text{ if } y_i = 1 \\ \frac{T_b(1-T_b)}{(T_a-T_b)^2} \text{ if } y_i = 0. \end{cases} \end{aligned}$$

2.3. Mangat's (1992) device

In Mangat's (1992) two-stage RR device each selected individual is requested to divulge truthfully, with a probability $q, (0 < q < 1)$, whether he/she bears the sensitive attribute A or the complementary attribute A^C and with probability $(1 - q)$ to report the RR obtained on applying Horvitz *et al.*'s (1967) or Greenberg *et al.*'s (1969) device in the following manner:
From a pack of cards marked A or B, B being a non-sensitive character in proportion $p : (1 - p), 0 < p < 1$, the sampled person is requested to draw a card, unnoticed by the interviewer and report truthfully whether the card type chosen 'matches' or 'mismatches' his/her true A or B characteristics.
Let us now define

$$x_i = \begin{cases} 1 \text{ if the } i\text{th individual bears } B \\ 0 \text{ if he/she bears } B^C \text{, the complement of } B. \end{cases}$$

Now since the values of x_i's usually will be unknown to the interviewer, to develop an estimator of π_A using Mangat's (1992) device Chaudhuri (2001a) prescribed the following solution.
Suppose that keeping all other things unchanged in Mangat's (1992) device, let there be two packs containing cards marked A and B in proportions $p_1 : (1 - p_1)$

and $p_2 : (1-p_2), (0 < p_1 < 1, 0 < p_2 < 1, p_1 \neq p_2)$ instead of only one pack of cards and a sampled person is requested to make two independent draws from the first pack and another two from the second and to report in each case whether his/her true A or B characteristic matches or mismatches with the card type. Now we define

$$I_i = \begin{cases} 1 & \text{if there is a match between the card type chosen from the} \\ & \text{first pack and the respondents true } A \text{ or } B \text{ character} \\ 0 & \text{else,} \end{cases}$$

$$J_i = \begin{cases} 1 & \text{if there is a match between the card type chosen from the} \\ & \text{second pack and the respondents true } A \text{ or } B \text{ character} \\ 0 & \text{else,} \end{cases}$$

and I_i', J_i' , are defined similarly to I_i and J_i respectively. Then it follows that

$$\begin{aligned} P(I_i = 1) &= qy_i + (1-q)[py_i + (1-p)x_i] = E_R(I_i) \\ &= E_R(I_i') \\ P(J_i = 1) &= qy_i + (1-q)[py_i + (1-p)x_i] = E_R(J_i) \\ &= E_R(J_i'). \end{aligned}$$

Now writing

$$r_{1i} = [(1-p_2)I_i - (1-p_1)J_i]/(p_1 - p_2) \text{ and } r_{2i} = [(1-p_2)I_i' - (1-p_1)J_i']/(p_1 - p_2)$$

we see that

$$E_R(r_{1i}) = y_i = E_R(r_{2i}) \ \forall i \in U$$

and also it follows that

$$V_R(I_i) = E_R(I_i)[1 - E_R(I_i)] = V_R(I_i') \text{ and } V_R(J_i) = E_R(I_i)[1 - E_R(I_i)] = V_R(J_i').$$

Again

$$V_R(r_{1i}) = \frac{(1-p_2)^2 V_R(I_i) + (1-p_1)^2 V_R(J_i)}{(p_1 - p_2)^2}$$

and

$$V_R(r_{2i}) = \frac{(1-p_2)^2 V_R(I_i') + (1-p_1)^2 V_R(J_i')}{(p_1 - p_2)^2}.$$

Thus, writing $r_i = (r_{1i} + r_{2i})/2$ it follows that $E_R(r_i) = y_i, \forall i \in U$ and

$$
\begin{aligned}
V_R(r_i) &= \frac{1}{4}[V_R(r_{1i}) + V_R(r_{2i})] \\
&= \frac{1}{2(p_1 - p_2)^2}[(1-p_2)^2 V_R(I_i) + (1-p_1)^2 V_R(J_i)] \\
&= \frac{(1-q)(y_i - x_i)^2(1-p_1)(1-p_2)}{2(p_1-p_2)^2}[p_1(1-p_2) + p_2(1-p_1) + 2q(1-p_1)(1-p_2)]
\end{aligned}
$$

An unbiased estimator of V_i is given by $v_i = (r_{1i} - r_{2i})^2/4$.

2.4. Mangat's (1994) device

Here each individual is instructed to respond 'yes' if he or she bears the sensitive attribute A. If he/she does not possess the character A, the respondent is requested to use Warner's (1965) device as defined in 2.1. Now defining

$$
I_i = \begin{cases} 1 \text{ if the } i\text{th selected individual responds 'yes'} \\ 0 \text{ else,} \end{cases}
$$

we have $P(I_i = y_i) = y_i + (1-p)(1-y_i)$. Then

$$
E_R(r_i) = y_i + (1-p)(1-y_i) = py_i + (1-p)P(I_i = y_i)
$$

and

$$
V_R(r_i) = E_R(I_i)[1 - E_R(I_i)] = p(1-p)(1-y_i)^2.
$$

Thus for $r_i = [I_i - (1-p)]/p, 0 < p < 1$ we have

$$
E_R(r_i) = y_i, \forall i \in U
$$

and

$$
\begin{aligned}
V_i = V_R(r_i) &= \frac{V_R(I_i)}{p^2} = \frac{1-p}{p}(1-y_i)^2 \\
&= \begin{cases} 0 & \text{if } y_i = 1 \\ \frac{(1-p)}{p} & \text{if } y_i = 0. \end{cases}
\end{aligned}
$$

Let E_p and V_p be the operators for expectation and variance respectively with respect to any arbitrary sampling design p and E, V be the overall expectation and variance operators. Now we shall write $E = E_p E_R = E_R E_p$. Suppose that $t = \sum_{i=1}^N y_i b_{si} I_{si}$ where $I_{si} = 1(0)$ if $i \in (\notin) s$, $i = (1, \cdots, i, \cdots, N)$ and b_{si}'s are constants free of $\underline{Y} = (y_1, \cdots, y_i, \cdots, y_N)$ and $\underline{R} = (r_1, \cdots, r_i, \cdots, r_N)$ such that $E_p(b_{si}I_{si}) = 1, \forall i \in U$ be an unbiased estimator for Y. We write

$$V_p(t) = \sum_{i=1}^{N} y_i^2 c_i + \sum_{\substack{j=1 \\ i \neq j}}^{N} y_i y_j c_{ij}$$

where $c_i = E_p(b_{si}^2 I_{si}) - 1, c_{ij} = E_p(b_{si} b_{sj} I_{sij}) - 1$ and $I_{sij} = I_{si} I_{sj}$.Then an unbiased estimator $v_p(t)$ of $V_p(t)$ is given by

$$v_p(t) = \sum_{i=1}^{N} y_i^2 c_{si} I_{si} + \sum_{\substack{j=1 \\ i \neq j}}^{N} y_i y_j c_{sij} I_{sij}$$

where c_{si}'s and c_{sij}'s are constants free of $\underline{Y}$, $\underline{R}$ such that $E_p(c_{si}I_{si}) = c_i$ and $E_p(c_{sij}I_{sij}) = c_{ij}$. Since y_i's are unknown and r_i's are unbiased for y_i's, an unbiased estimator for Y is obtained as

$$e = \sum_{i=1}^{N} r_i b_{si} I_{si}.$$

Note that

$$\begin{aligned} V(e) = E_p E_R (e - Y)^2 &= E_p E_R (e - t + t - Y)^2 = V_p(t) + E_p E_R (e - t)^2 \\ &= V_p(t) + E_p[\sum_{i=1}^{N} b_{si}^2 I_{si} V_i] \end{aligned}$$

and

$$V(e) = E_R E_p (e - Y)^2 = E_R V_p(e) + E_R (R - Y)^2 = E_R V_p(e) + \sum_{i=1}^{N} V_i$$

where $R = (\sum_{i=1}^{N} r_i)/N$. Now taking v_i as an unbiased estimator for $V_i = V_R(r_i)$ and following Raj (1968) and Rao (1975)two unbiased estimators for $V(e)$ are given by

$$v_1(e) = v_p(t) \mid_{\underline{Y}=\underline{R}} + \sum_{j=1}^{N} b_{si} I_{si} v_i \tag{2.1}$$

$$v_2(e) = v_p(t) \mid_{\underline{Y}=\underline{R}} + \sum_{j=1}^{N} (b_{si}^2 - c_{si}) I_{si} v_i. \tag{2.2}$$

3. PROTECTION OF PRIVACY AND THE RELATED MEASURES OF JEOPARDY

Let Ω be the set of all possible alternative forms in which an actual RR may turn out for a given device. For the devices discussed here the RR's are "yes" or "no" depending on whether the card type or outcome of the random experiment matches or mismatches the individual's true A or A^C- characteristic. Then writing

$$P(yes|A) = \alpha \text{ and } P(no|A^C) = \beta$$

for any RR scheme based on SRSWR, Nayak (1994) noted the revealing posterior probabilities on applying Bayes' theorem as

$$P(A|yes) = \frac{\alpha\pi_A}{\alpha\pi_A + (1-\beta)(1-\pi_A)}, \quad P(A|no) = \frac{(1-\alpha)\pi_A}{(1-\alpha)\pi_A + \beta(1-\pi_A)}$$

since $P(A) = \pi_A$ and $P(A^C) = 1 - \pi_A$.

In general, if R be a response "yes" or "no" then the conditional probabilities that the observation is in A or A^C given the response R are

$$\begin{aligned} P(A|R) &= \frac{\pi_A P(R|A)}{\pi_A P(R|A) + (1-\pi_A)P(R|A^C)} \\ P(A^C|R) &= \frac{(1-\pi_A)P(R|A^C)}{(1-\pi_A)P(R|A^C) + \pi_A P(R|A)}. \end{aligned}$$

Now the response R is said to be jeopardizing with respect to A or A^C according as

$$P(A|R) > \pi_A \quad \text{or} \quad P(A^C|R) > 1 - \pi_A.$$

Since $P(R|A)$ and $P(R|A^C)$ are the design probabilities of R for an observation in A or A^C, a measure of jeopardy based on $P(R|A)$ and $P(R|A^C)$ is defined by

$$JM_1(R) = \frac{P(R|A)}{P(R|A^C)} = \frac{P(A|R)/\pi_A}{P(A^C|R)/(1-\pi_A)}. \tag{3.1}$$

If R is only a "yes" or "no" response variable, (3) is known as the jeopardy of R with respect to A and the jeopardy of R with respect to A^C is defined as

$$JM_2(R) = \frac{P(R|A^C)}{P(R|A)} = \frac{P(A^C|R)/(1-\pi_A)}{P(A|R)/\pi_A}. \tag{3.2}$$

Note that $JM_1(R)$ gives the measure of jeopardy of the response R with respect to A and $JM_2(R)$ is the jeopardy of R with respect to A^C. Here both $JM_1(R)$

and $JM_2(R)$ depend on the randomized response R and these measures can not be treated as an overall measure of jeopardy of the RR device. If a measure of jeopardy of an RR device is to be developed, it has to be independent of the RR provided by the interviewees. Thus taking the simple average of $JM_1(R)$ and $JM_2(R)$ a new measure, say, JM independent of the response R can be defined as

$$JM = \frac{1}{2}[JM_1(R) + JM_2(R)] \tag{3.3}$$

and this measure can be regarded as a measure of jeopardy of the RR device.

In order to extend these concepts to more general set-up where the respondents are selected with unequal probabilities, we suppose that $L_i = P(y_i = 1), 0 < L_i < 1$, *i.e.*, L_i is the prior probability, $P(y_i = 1)$ assigned to the ith individual and $L_i(R) = P(y_i = 1|\text{the response is } R)$, *i.e.*, $L_i(R)$ be the posterior probability that the ith respondent bears the sensitive characteristic A when the RR offered is 'R'.

Then the response-specific jeopardy for the RR obtained as R is given by

$$JM(R) = \frac{P(R|A)}{P(R|A^C)} = \frac{L_i(R)/L_i}{[1 - L_i(R)]/(1 - L_i)}. \tag{3.4}$$

One may note that for any RR device R may take more than one value, *e.g.*, for the RR devices discussed here R takes two values, viz., '1' and '0' and considering this a measure of jeopardy independent of R can be defined as

$$JM = \frac{1}{\eta}\sum_{R\in\Omega} JM(R) \tag{3.5}$$

where η is the cardinality of Ω.

Now the expressions for JM and $L_i(1)$ for the four RR-devices discussed above are given below:

(i) Warner's(1965) device

$$\begin{aligned} L_i(1) &= \frac{L_iP(I_i = 1|y_i = 1)}{L_iP(I_i = 1|y_i = 1) + (1 - L_i)P(I_i = 1|y_i = 0)} \\ &= \frac{pL_i}{(1-p) + (2p-1)L_i}; \end{aligned}$$

$JM(1) = p/(1-p), JM(0) = (1-p)/p$ and $JM = [JM(1) + JM(0)]/2$. Note that as $p \to 1/2, L_i(1) \to L_i, JM(1) \to 1, JM(0) \to 1$ and $JM \to 1$, but $V_i \to \infty$ as $p \to 1/2$. Thus we see that the better is the protection to the respondents the worse is the estimator.

(ii) Chaudhuri and Mukerjee's (1988) Modification of Warner's device

$$\begin{aligned} L_i(1) &= \frac{L_iP(I_i=1|y_i=1)}{L_iP(I_i=1|y_i=1)+(1-L_i)P(I_i=1|y_i=0)} \\ &= \frac{L_iT_a}{L_iT_a+(1-L_i)T_b} \to L_i \end{aligned}$$

as $T_a \to T_b$. Note that $L_i(1) \to L_i$ as $T_a \to T_b$ and $JM(1) \to 1, JM(0) \to 1, JM \to 1$ as $T_a \to T_b$. But $V_i \to \infty$ as $T_a \to T_b$ implying that better protection of privacy results in less efficient estimators.

(iii) Mangat's (1992) Scheme
Here

$$\begin{aligned} L_i(1,1) &= \frac{L_iP(I_i=1,J_i=1|y_i=1)}{L_iP(I_i=1,J_i=1|y_i=1)+(1-L_i)P(I_i=1,J_i=1|y_i=0)} \\ &= \frac{L_i[q+(1-q)p_1][q+(1-q)p_2]}{L_i[q+(1-q)p_1][q+(1-q)p_2]+(1-L_i)(1-p_1)(1-p_2)} \\ &\to \frac{L_ip_1p_2}{(1-p_1)(1-p_2)+L_i(p_1+p_2-1)} \text{ as } q\to 0 \\ &\to L_i \text{ as } p_1 \to 1/2, p_2 \to 1/2 \end{aligned}$$

and

$$\begin{aligned} JM(1,1) &= \frac{[q+(1-q)p_1][q+(1-q)p_2]}{(1-q)(1-p_1)(1-p_2)}, JM(1,0) = \frac{[q+(1-q)p_1](1-q)(1-p_2)}{(1-q)(1-p_1)p_2}, \\ JM(0,1) &= \frac{[q+(1-q)p_2](1-q)(1-p_1)}{(1-q)(1-p_2)p_1}, JM(0,0) = \frac{(1-p_1)(1-p_2)}{[q+(1-q)p_1][q+(1-q)p_2]}. \end{aligned}$$

Thus $JM = [JM(1,1)+JM(1,0)+JM(0,1)+JM(0,0)]/4$ and note that $JM(1,1) \to 1, JM(1,0) \to 1, JM(0,1) \to 1, JM(0,0) \to 1$ and $JM \to 1$ as $q \to 0$, $p_1 \to 1/2, p_2 \to 1/2$. But for such choices of p_1 and p_2, V_i becomes unbounded.

(iv) Mangat's (1994) Scheme
Here

$$\begin{aligned} L_i(1,1) &= \frac{L_iP(I_i=1|y_i=1)}{L_iP(I_i=1|y_i=1)+(1-L_i)P(I_i=1|y_i=0)} \\ &= \frac{L_i}{L_i+(1-p)(1-L_i)} \\ &\to L_i \text{ as } p \to 0. \end{aligned}$$

TABLE 4.1 *Showing the values of $L_i(1), V_i, JM(1), JM(0)$ for different choices of design parameters Warner's device*

$L_i \rightarrow$	0.10	0.20	0.30	0.40	0.50	0.60	0.70	V_i	measures of jeopardy		
p				$L_i(1)$					$JM(1)$	$JM(0)$	JM
0.51	0.104	0.206	0.308	0.410	0.510	0.610	0.708	624.750	1.040	0.961	1.001
0.47	0.090	0.181	0.275	0.372	0.470	0.571	0.674	69.194	0.887	1.128	1.007
0.55	0.120	0.234	0.344	0.449	0.550	0.647	0.740	24.750	1.222	0.818	1.020
0.39	0.066	0.138	0.215	0.299	0.390	0.490	0.599	4.915	0.639	1.564	1.102
0.67	0.184	0.337	0.465	0.575	0.670	0.753	0.826	1.913	2.030	0.493	1.261

TABLE 4.2 *Showing the values of $L_i(1), V_i, JM(1), JM(0)$ and JM for different choices of design parameters Chaudhuri and Mukerjee's Modification of Warner's device*

$L_i \rightarrow$		0.10	0.20	0.30	0.40	0.50	0.60	0.70	V_i		measures of jeopardy		
T_a	T_b				$L_i(1)$				$y_i = 1$	$y_i = 0$	$JM(1)$	$JM(0)$	JM
0.18	0.50	0.038	0.083	0.134	0.194	0.265	0.351	0.457	3.441	2.441	0.360	1.640	1.000
0.70	0.77	0.092	0.185	0.280	0.377	0.476	0.577	0.680	29.429	36.143	0.909	1.304	1.107
0.36	0.24	0.143	0.273	0.391	0.500	0.600	0.692	0.778	9.333	12.667	1.500	0.842	1.171
0.76	0.84	0.091	0.184	0.279	0.376	0.475	0.576	0.679	13.500	21.000	0.905	1.500	1.202
0.76	0.24	0.260	0.442	0.576	0.679	0.760	0.826	0.881	0.675	0.675	3.167	0.316	1.741

Also $JM(1) = 1/(1-p), JM(0) = 1-p$ and $JM = 1/(1-p) + (1-p)$. Thus we see that for this device $JM(1) \rightarrow 1, JM(0) \rightarrow 1$ and $JM \rightarrow 1$ as $p \rightarrow 0$ while $V_i \rightarrow \infty$ as $p \rightarrow 0$. Hence for Mangat's (1994) device also protection of privacy and measure of jeopardy move in opposite directions.

4. NUMERICAL ILLUSTRATIONS

We now present below some numerical results illustrating the values of $L_i(1)$, $L_i(1,1)$, V_i and the measures of jeopardy for different values of the design parameters for the four methods we discussed in section 2.

It is observed from Table 4.1 that as $p \rightarrow 1/2$, $L_i(1) \rightarrow L_i$, $JM(1), JM(0)$ and $JM \rightarrow 1$, but for such choices of p, $V_i \rightarrow \infty$. This implies that the better is the protection to the respondents the worse is the performance of the estimator.

As in the case of Warner's device the efficacy of the estimator decreases with increased privacy protection for Chaudhuri and Mukerjee's modification and for Mangat's methods also. This is evident from Tables 4.2, 4.3 and 4.4 above. In Table 4.3, as $q \rightarrow 0$, $p_1 \rightarrow 1/2, p_2 \rightarrow 1/2, L_i(1,1) \rightarrow L_i, JM(1,1) \rightarrow 1, JM(1,0) \rightarrow 1, JM(0,1) \rightarrow 1, JM(0,0) \rightarrow 1$ and $JM \rightarrow 1$, but for such choices of p_1 and p_2, $V_i \rightarrow \infty$. Similar pattern is also observed for Chaudhuri and Mukerjee's modifi-

TABLE 4.3 *Showing the values of $L_i(1,1), V_i, JM(1,1), JM(1,0), JM(0,1), JM(0,0)$ and JM for different choices of design parameters Mangat's (1992) device for $q = 0.05$*

$L_i \rightarrow$		0.20	0.40	0.60	0.80	V_i	measures of jeopardy				
p_1	p_2	$L_i(1,1)$					$JM(1,1)$	$JM(1,0)$	$JM(0,1)$	$JM(0,0)$	JM
0.45	0.51	0.191	0.392	0.591	0.790	70.830	0.961	0.890	1.130	1.040	1.000
0.53	0.49	0.211	0.419	0.620	0.812	37.501	1.080	1.171	0.850	0.919	1.010
0.35	0.47	0.111	0.240	0.424	0.661	5.870	0.479	0.611	1.650	2.091	1.210
0.69	0.28	0.182	0.373	0.563	0.782	0.392	0.870	5.719	0.171	1.160	1.980
0.30	0.61	0.142	0.311	0.500	0.733	0.770	0.668	0.272	3.646	1.494	1.520

TABLE 4.4 *Showing the values of $L_i(1), V_i, JM(1), JM(0)$ and JM for different choices of design parameters Mangat's(1994) device*

$L_i \rightarrow$	0.10	0.20	0.30	0.40	0.50	0.60	0.70	V_i	measures of jeopardy		
p	L_i								$JM(1)$	$JM(0)$	JM
0.05	0.105	0.208	0.311	0.412	0.513	0.612	0.711	19.000	1.053	0.950	1.001
0.10	0.110	0.217	0.322	0.425	0.526	0.625	0.722	9.000	1.111	0.900	1.005
0.15	0.116	0.227	0.335	0.439	0.540	0.638	0.733	5.667	1.176	0.850	1.013
0.20	0.122	0.238	0.349	0.454	0.555	0.652	0.745	4.000	1.250	0.800	1.025
0.40	0.156	0.294	0.366	0.526	0.625	0.714	0.795	1.500	1.667	0.600	1.133

cation and also for Mangat's (1994) method.
Here from the illustrations in Tables 4.1 to 4.4, it is may be noted that once the experimenter chooses the value(s) of the design parameter(s), the measure of respondent's jeopardy as given by our new response independent measure, JM can be calculated well ahead of conducting the experiment. And if required, the respondent may be informed of the jeopardy before he/she agrees to use the RR device for providing the RR. This may help in ensuring better respondent cooperation.

Thus summarizing from the tables above, it is observed that with an increase in the protection level of respondent's privacy the efficiency of the estimator of the population parameter decreases and it is almost impossible to retain respondent's jeopardy at an ideal level keeping the efficiency level of the estimator at a reasonably high limit. So for choosing the design parameters of an RR device, one has to strike a balance between the respondent's jeopardy and efficiency of the estimator of the population parameter. Of course, this is true for the RR devices discussed here and there may exist other RR devices for which it may be possible to keep both the jeopardy and efficiency level at desired limits. A general tendency of the researchers and experimenters is to choose the design parameter(s) in such a manner that it produces efficient estimator(s) of the population parameter(s). But as discussed above, often such choice(s) of the design parameter(s) jeopardize the respondents' privacy. Our new measure of jeopardy will provide a guideline to the experimenter in choosing the value(s) of the design parameter(s) for keeping the level of respondents' jeopardy as well as efficacy of the estimator(s) within reasonable limits.

ACKNOWLEDGEMENT

The author gratefully acknowledges the help and guidance received from Prof. Arijit Chaudhuri of the Indian Statistical Institute. Helpful comments and recommendations from two unanimous referees is also sincerely acknowledged.

REFERENCES

ANDERSON, H. (1977). "Efficiency versus protection in a general RR-model", *Scandinavian Journal of Statistics*, **4**, 11–19.

ANDERSON, H. (1976). "Estimation of a Proportion through Randomized Response", *International Statistical Review*, **44**(2), 213–217.

CHAUDHURI, A. (2001a). "Using randomized response from a complex survey to estimate a sensitive proportion in a dichotomous finite population", *Journal of Statistical Planning and Inference*, **94**, 37–42.

CHAUDHURI, A. (2001b). "Estimating sensitive proportions from unequal probability sample using randomized responses", *Pakistan Journal Statistics*, **17**(3), 259–270.

CHAUDHURI, A. AND MUKERJEE, R. (1988). *Randomized response: theory and techniques*, Marcel Dekker Inc. New York.

GREENBERG, B.G., ABUL-ELA, SIMMONS, W.R. AND HORVITZ, D.G. (1969). "The unrelated question randomized response model: theoretical framework", *Journal of the American Statistical Association*, **64**, 520–539.

HORVITZ, D.G., SHAH, B.V. AND SIMMONS, W.R. (1967). "The unrelated question randomized response model", *Proc. Soc. Sect. Amer. Statis. Assoc.*, 65–72.

LANKE, J. (1976). "On the degree of protection in randomized interviews", *International Statistical Review*, **44**, 197–203.

LEYSIEFFER, R. W. AND WARNER, S. L. (1976). "Respondent jeopardy and optimal designs in RR models", *Journal of the American Statistical Association*, **71**, 649–656.

MANGAT, N. S. (1994). "An improved randomized response strategy", *Journal of the Royal Statistical Society*, **B56**, 93–95.

MANGAT, N. S. (1992). "Two stage randomized response sampling procedure using unrelated selection", *Journal of Indian Society of Agricultural Statistics*, **44**(1), 82–87.

NAYAK, TAPAN K. (1994). "On randomized response survey for estimating a proportion", *Communications in Statistics: Theory and Methods*, **23**(11), 3303–3321.

RAJ, DES (1968). *Sampling Theory*, Mc-Graw Hill, New York.

RAO, J.N.K (1975). "Unbiased variance estimation for multi-stage designs", *Sankhya C, 37*, 133–139.

WARNER, S.L. (1965). "Randomised response: a survey technique for eliminating evasive answer bias", *Journal of the American Statistical Association*, **60**, 63–69.